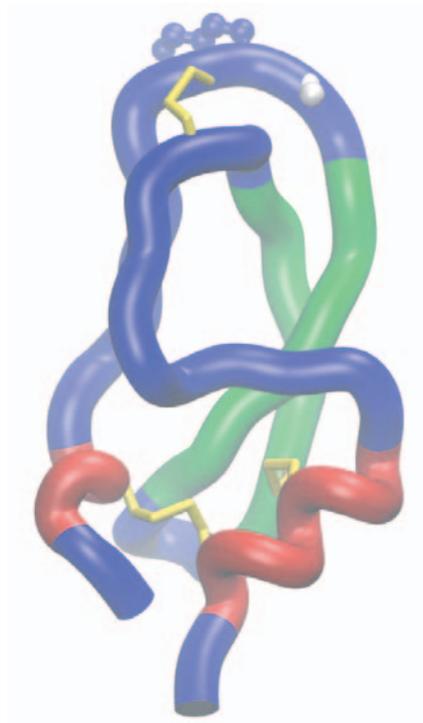


Case Study: BPTI

Markus Dittrich and Chalermpol Kanchanawarin



1 Introduction

Bovine pancreatic trypsin inhibitor (BPTI) is one of the smallest and simplest globular proteins. BPTI's function is the suppression of protein digestion, i.e., the breakdown of proteins into their peptide building blocks, by means of inhibiting the action of the enzyme trypsin which is produced in the bovine pancreas. BPTI is a member of the serine protease family of inhibitors [1]. A hallmark of this class of enzymes are the many conserved cysteine residues that form disulfide bonds stabilizing the proteins' three-dimensional structures. BPTI has a relatively broad specificity in that it can inhibit several kinds of digestive enzymes [1].

BPTI is one of the most extensively studied globular proteins and has been investigated structurally by both x-ray crystallography and NMR spectroscopy. Furthermore, its protein folding pathway and dynamics have been investigated in great detail. Since BPTI forms complexes with several of the enzymes that it inhibits, it has also been the subject of studies investigating protein-protein interaction and molecular recognition [1].

This case study is divided into three sections: First, the primary sequence and the three-dimensional structure of BPTI will be introduced. The second part will examine the thermal stability of BPTI that is caused by its three internal disulfide bonds. Finally, in the last section, we study BPTI's function as trypsin inhibitor by investigating a trypsin-BPTI complex.

Figure 1 shows the files required for the Exercises in the present case study. The files can be found in the directory

`BPTI/00-FORUSERS/supplementary-files/`.

2 Sequence and structure of BPTI

BPTI is a single-chain polypeptide that contains 58 amino acid residues and has a molecular mass of 6512 [1]. Atomic resolution X-ray crystal structures of BPTI and several of its mutants have been available since 1970 [1, 3].

Fig. 2(a) shows the three dimensional conformation of BPTI and the image reveals that BPTI has a compact pear-shape with a maximum dimension of about 30Å. Topologically, the protein folds like a piece of string that is folded upon itself twice as shown in Figure 2(b). Three disulfide bonds, D1, D2 and D3, link the protein's secondary structure elements together.

The secondary structure of BPTI is shown in Fig. 3 along with its amino acid sequence. Two α helical regions, H1 and H2, form a one and a half turn 3_{10} -helix near the N-terminus and a three turn α -helix near the C-terminus. The central part of the protein contains two β -strands, B1 and B2, that are connected by a turn, T1, and which form an anti-parallel β -hairpin loop. As already mentioned, six conserved cysteine residues in BPTI form three disulfide bonds that stabilize the three dimensional structure of the protein. Cys5 and Cys55 form the first disulfide bond D1 that joins the N- and C-termini of BPTI. Cys14 and Cys38 make the second disulfide bond

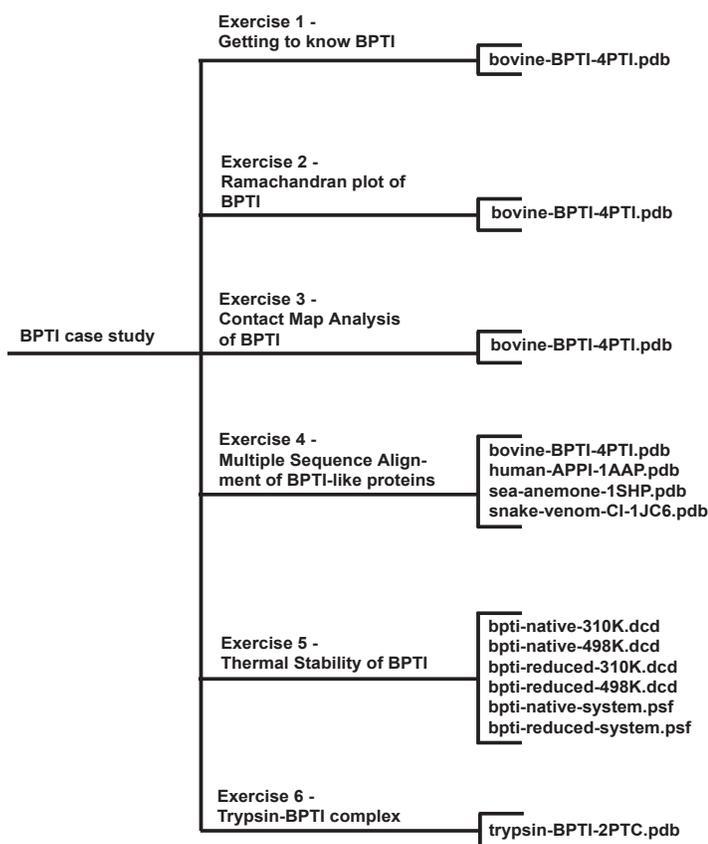


Figure 1: Required files for exercises.

D2 which connects the loose end of the two β -strands. The third disulfide bond, D3, formed by Cys30 and Cys51 helps to pack the three turn α -helix against the double stranded anti parallel β sheet.

Exercise 1: Getting to know BPTI

After the previous brief introduction to BPTI, we will use VMD to investigate its structure in more detail. First, load the BPTI coordinate file `bpti/bovine-BPTI-4PTI.pdb` into VMD and then select the rendering options from the VMD *Graphical Representations Tool* such that the structure looks roughly similar to what

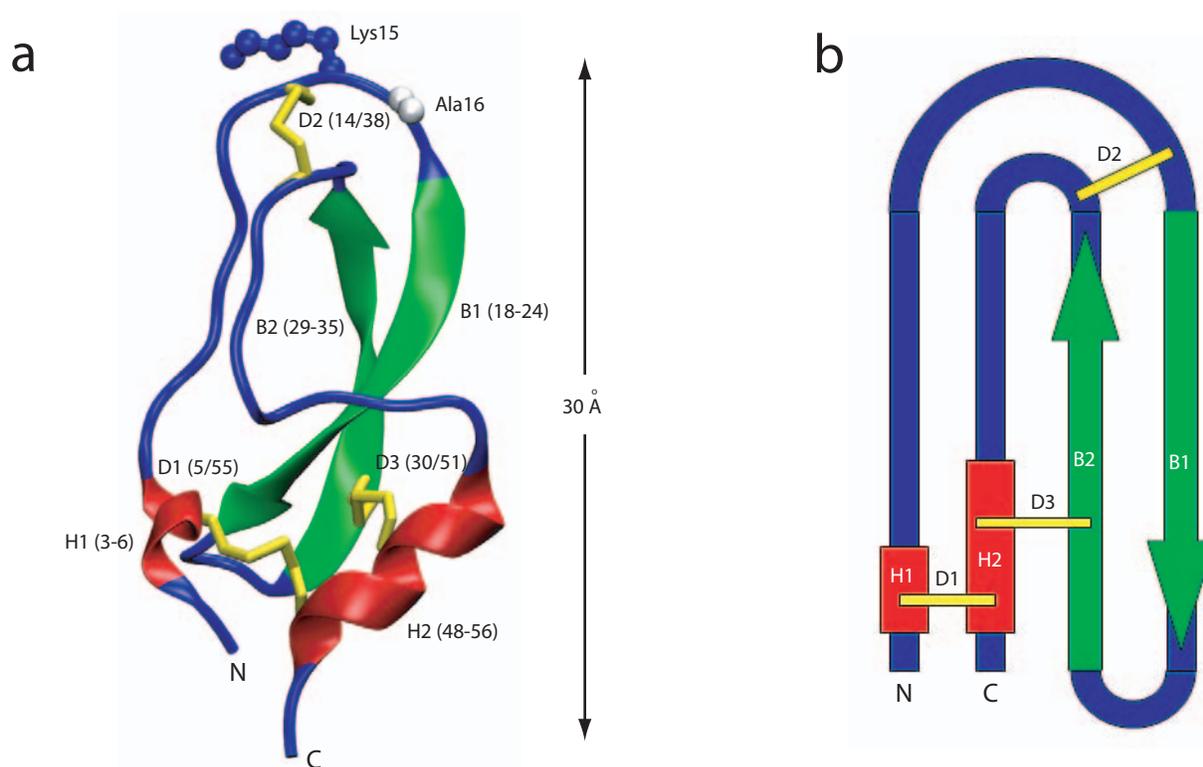


Figure 2: Tertiary structure of BPTI. (a) The structure of BPTI. Helices are shown in red, β strands are shown in green, loops and turns are drawn in blue, and disulfide bonds are shown in yellow. The structure is taken from the Protein Data Bank with PDB ID: 4PTI. (b) Folding topology of BPTI. *Self-study:* Using the pdb file 4PTI and VMD, try to reproduce (a), in particular the location of the three disulfur bonds.

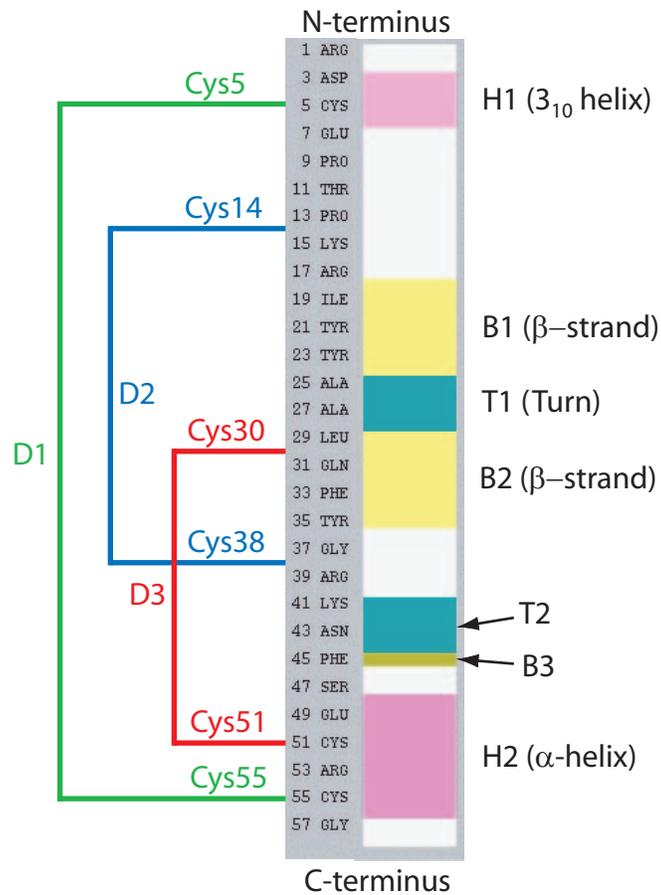


Figure 3: Primary sequence and secondary structure of BPTI. The diagram shows the amino acid sequence of BPTI (only odd numbered residues are labeled explicitly). The secondary structures are represented using the following color codes: α -helix = purple, 3_{10} helix = pink, β -strand = yellow, turn = green, partial β -strand = gold, coil = white. Also shown are the three disulfur bonds D1, D2, and D3 that link BPTI's secondary structure elements. *Self-study:* Using the the pdb file 4PTI and VMD's sequence browser try to identify the disulfur bonds and secondary structure elements shown in Fig. 2 and, thereby, connect the primary, secondary, and tertiary structure of BPTI.

is shown in Fig. 2 (a). Now use VMD to highlight the following parts of the BPTI structure using the *Sequence Viewer* and the

Graphical Representations Tool:

- The one and a half turn 3_{10} helix (Asp3 to Leu6) and the three turn α -helix (Ala48 to Gly56).
- The two β -strands forming an anti-parallel β -sheet (Ile18 to Asn24 and Leu29 to Try35).
- The three disulfide bonds D1, D2, and D3 (Cys5/55, Cys14/38 and Cys30/51).
- The reactive site which binds trypsin or other proteolytic enzymes, located between residues Lys15 and Ala16, near the disulfide bond D2 (Cys14/38).

More challenging, try to identify the following more detailed structural features:

- Two β turns (Ala25 to Gly28 and Lys41 to Asn44).
- A partial third β -strand formed by Phe45 and Tyr21. The key observation to make here is the hydrogen-bonding pattern of the backbone oxygen and amine hydrogen atoms in β -sheets.
- The hydrophobic core of BPTI consisting of Phe4, Cys5/55, Phe22, Tyr23, Cys30/51, Phe33, Tyr35 and Phe45. What is its significance in terms of the folding and overall stability of BPTI's structure?
- The three hydrogen bonds formed by Asn43 and the amine hydrogen and carbonyl oxygen of Tyr23 and the carbonyl oxygen of Gly7. What possible roles could these hydrogen bonds have?

A very convenient way to quickly obtain an overview over the secondary structure elements present in a protein is to simultaneously plot the (ϕ, ψ) angle pairs of the peptide bond in a so called Ramachandran plot as shown in Fig. 4 for BPTI. Here, each pair of (ϕ, ψ) angles (corresponding to a single amino acid residue) is plotted in a two dimensional plane spanned by the ϕ and ψ -axis whose values go from -180° to 180° , respectively. The blue areas indicate allowed conformations and correspond to β -sheets in the fourth quadrant (labeled β in Fig. 4) and α -helices in the third quadrant (labeled

α in Fig. 4). The green areas represent partially allowed conformations. All other pairs of (ϕ, ψ) angles are typically absent from protein structures. As illustrated nicely in Fig. 4, BPTI is no exception to this rule, and most (ϕ, ψ) angles actually fall into the allowed regions.

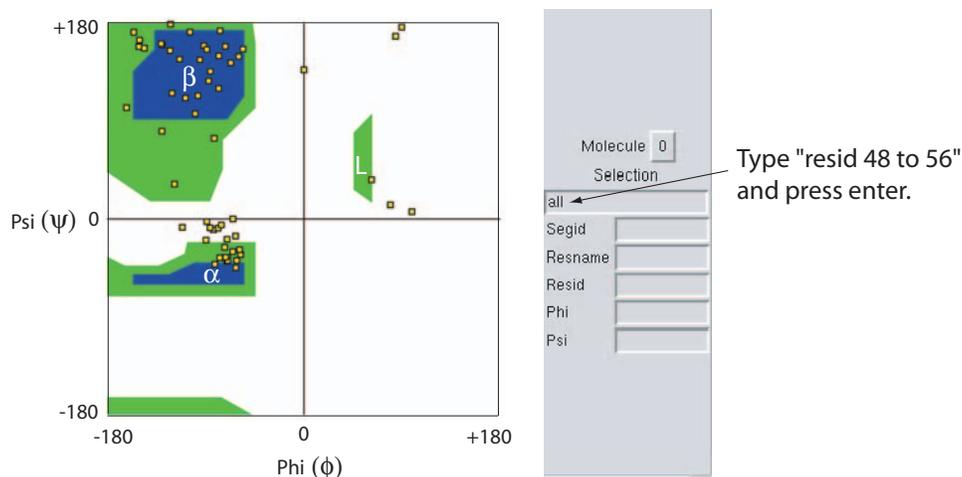


Figure 4: Ramachandran plot of BPTI. The figure shows the pairs of (ϕ, ψ) angles for all 58 amino acid residues of BPTI and was created using the *Ramachandran Plot Tool* of VMD. The blue and green colored regions marked α , β and L indicate (ϕ, ψ) values corresponding to right handed α -helices, β -strands and left handed helices. Here, regions colored blue indicate the most allowed angles and green the partially allowed angles.

Exercise 2: Ramachandran plot of BPTI

As a continuation of Exercise 1, we will now use VMD's *Ramachandran Plot Tool* to study BPTI's structural properties in more detail.

Again, the BPTI coordinate file `bpti/bovine-BPTI-4PTI.pdb` will be used as starting point. Hence, load the structure into VMD, open up the *Ramachandran Plot Tool* and consider the following tasks:

- Type “resid 48 to 56” in the selection box. Do these residues form an α -helix?

- Type “all” in the selection box. There are a few dots outside the blue and green regions. Find out what residues they correspond to by clicking on them and explain your findings!

In addition to a Ramachandran plot, a so called contact map consisting of a two dimensional plot of distances between the C_α atoms of any pair of amino acid residues in a protein can provide important and useful information regarding its the tertiary structure. Fig. 5 shows the contact map generated for BPTI, in which the pairwise C_α distances are displayed in a gray scale ranging from 0 to 10Å. It is apparent, that the distance map immediately reveals the presence of the central β -sheet comprised of B1-B2.

Exercise 3: Contact Map Analysis of BPTI

In this exercise, we will use the *Contact Map Tool* of VMD to establish the spatial arrangement of secondary structure elements in BPTI. To this end, load the BPTI structure file `bpti/bovine-BPTI-4PTI.pdb` into VMD, open up the *Contact Map Tool*, and generate the map. With the help of the displayed contact map, identify the following contacts between parts of BPTI:

- Contact between β -strands B1 and B2 (c.f. Fig. 5).
- Contact between α -helix H1 and β -strand B2.
- Contact between the 3_{10} -helix and β -strand B1.

After having learned about the structural properties of BPTI, it is instructive to study other members of the so called BPTI-like superfamily. The latter is characterized by the presence of three disulfide bonds stabilizing their respective tertiary structures. All these proteins are structurally similar to BPTI but differ significantly in their amino acid sequences and, importantly, their physiological function.

Here, we consider three particular examples and compare them with BPTI. Our first example of a BPTI-like protein is the human BPTI-domain of the amyloid β -protein precursor [7] implicated in Alzheimer’s disease. The second one is a chymotrypsin inhibitor isolated from snake venom [5]. Finally, we consider a BPTI-like protein taken from a Caribbean sea anemone [6].

First, we will examine the backbone conformation of bovine BPTI and the other three BPTI-like proteins. They are depicted in Fig. 6(a) in tube

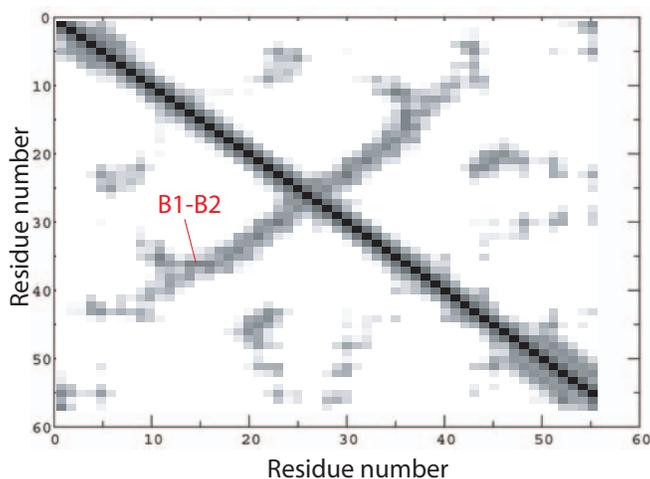


Figure 5: Contact map of BPTI. Shown are the distances between two C_{α} atoms of amino acids on BPTI in gray scale ranging from 0 to 10 Å (black to white).

representation. By inspection, it can be seen that the core region of all proteins are structurally very similar to that of BPTI. The main difference are the lengths and the conformation of the N and C-termini, most apparent in the case of the C-terminus from the snake venom protein, which is significantly elongated. This structural similarity becomes clearer after aligning all four protein structures using multiple sequence alignment. The so aligned structures are shown in Fig. 6(b) and reveal the proteins' common core conformation and their variable termini. Fig. 6(c) makes this even more apparent. It shows the same aligned structures but this time colored by Q-value which is a measure of structural similarity, blue and red indicating high and low similarity, respectively. Interestingly, this similarity comes despite the fact that the underlying sequence identity among all four structures is rather low as shown in Fig. 6(d), where blue and red indicate high and low sequence similarity, respectively. Fig. 6(e) highlights the conserved residues in the primary sequence that include the cysteines of the three conserved disulfur bridges which are the hallmark of the BPTI-like proteins. Hence, as this example clearly shows, different primary amino-acid sequences can give rise to very similar secondary and tertiary protein structures, a finding that has emerged over and over again when examining members of structurally closely related protein families.

proteins

Using VMD and its *Multiple Sequence Alignment Tool* [4] we will now take a closer look at the three BPTI-like proteins and then compare their 3D structures and amino acid sequences with that of BPTI. First, load the coordinates of the three BPTI-like proteins by opening the following PDB files: `human-APPI-1AAP.pdb`, `snake-venom-CI-1JC6.pdb` and `sea-anemone-1SHP.pdb` and then render all the proteins similar to the way they look in Fig. 6(a) using VMD's tube representation and coloring by segment name.

To allow for a better comparison of the conformational differences between the four proteins, we align them using the structure alignment option in the *Multiple Sequence Alignment Tool* as shown in Fig. 6 (b). Clearly, the 3D structures of the three BPTI-like proteins align very well with that of BPTI. The 3D structure of the BPTI-like protein from snake venom seems to deviate most from that of BPTI. Now, try the following:

- Color the four proteins according to the Q value (see Fig. 6 (c)) and find a part of one of the BPTI-like proteins (other than the N- or C- termini) that has its structure deviate most from the other proteins. What species is that BPTI-like protein from?
- Color the four BPTI-like proteins according to sequence identity and identify the residues that are highly conserved. What kinds of amino acids are they and where are they in the structures? Why do you think they are highly conserved?
- Use residue selection by Q value to find parts of the proteins' structures that are more than 80% conserved structurally. Where on the structures are they located?
- Use the multiple alignment tool to generate a phylogenetic tree of the four proteins according to the Q value. Can you tell which BPTI-like protein is closest to BPTI in terms of evolutionary history?

3 Structural and thermal stability of BPTI

Most globular proteins are stabilized by their hydrophobic core which is formed by the internal packing of non-polar amino acid side chains. BPTI is no exception and due to its rather small overall size it also has a relatively small hydrophobic core. Thus, extra stability to support its three dimensional structure is provided by the three internal disulfide bonds that were discussed previously.

It is, therefore, interesting to consider what would happen if any or several of these disulfide bonds are being cleaved. This allows one to judge how important the disulfide bonds are for preserving BPTI's three dimensional structure.

With all of its three disulfide bonds present, BPTI is one of the most stable globular proteins. Below 100°C, BPTI is quite inert to urea, a substance often used to denature proteins and only denatures under very acidic conditions. However, if a single one of the three disulfide bonds is reduced (cleaved), e.g. D2 (Cys14/Cys38), BPTI can be denatured quite easily. Fig. 7 shows the percentage of native BPTI that is denatured at various temperatures compared to reduced BPTI with a cleaved D2 disulfide bond. It can be seen that the native BPTI is quite stable as a function of temperature with only 50% denatured at about 85°C. On the other hand, when D2 is cleaved, BPTI becomes much less stable with 50% denatured at about 55°C. Finally, experimentally it is found that when all the three disulfide bonds are removed, BPTI unfolds at room temperature.

Exercise 5: Thermal stability of BPTI

We will now have a closer look at the thermal stability of native and fully reduced BPTI (i.e., all three disulfur bonds are in the cleaved state) by investigating several computer simulation trajectories. The latter provide a dynamic view of the time evolution of a biomolecular system. Four different trajectories of native and fully reduced BPTI immersed in water as solvent at two different temperatures will be considered. The respective lengths of the trajectories are given in nano seconds (ns, 10^{-9} s):

- T = 310K (0.7ns for the native system and 1.0ns for the reduced system) and
- T = 498K (2.3ns for the native system and 1.7ns for the reduced system).

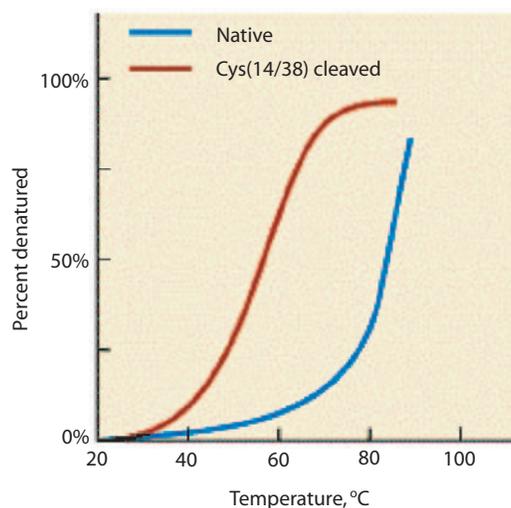


Figure 7: Temperature denaturation of native and single disulfide reduced BPTI.

Fig. 8 shows structural views of the systems for native BPTI at times $t=0$, $t=0.7\text{ns}$ ($T=310\text{K}$)/ $t=2.3\text{ns}$ ($T=498\text{K}$) and reduced BPTI at times $t=0$, $t=1.0\text{ns}$ ($T=310\text{K}$)/ $t=1.7\text{ns}$ ($T=498\text{K}$), respectively. To start the investigation, load the following four trajectory (dcd) files and protein structure (psf) files into VMD

- `bpti-native-310K.dcd` and `bpti-native-system.psf`
- `bpti-native-498K.dcd` and `bpti-native-system.psf`
- `bpti-reduced-310K.dcd` and `bpti-reduced-system.psf`
- `bpti-reduced-498K.dcd` and `bpti-reduced-system.psf`.

Note that each trajectory file shows trajectory frames only for every 25 ps to reduce the file size. Now, looking at the trajectories of native BPTI at $T=310\text{K}$ and $T=498\text{K}$ try to answer the following question:

- From looking at the trajectories, is native BPTI stable at $T=310\text{K}$ or $T=498\text{K}$?

Next, have a look at the trajectories of native BPTI and the reduced BPTI at $T=498\text{K}$ and consider the following questions:

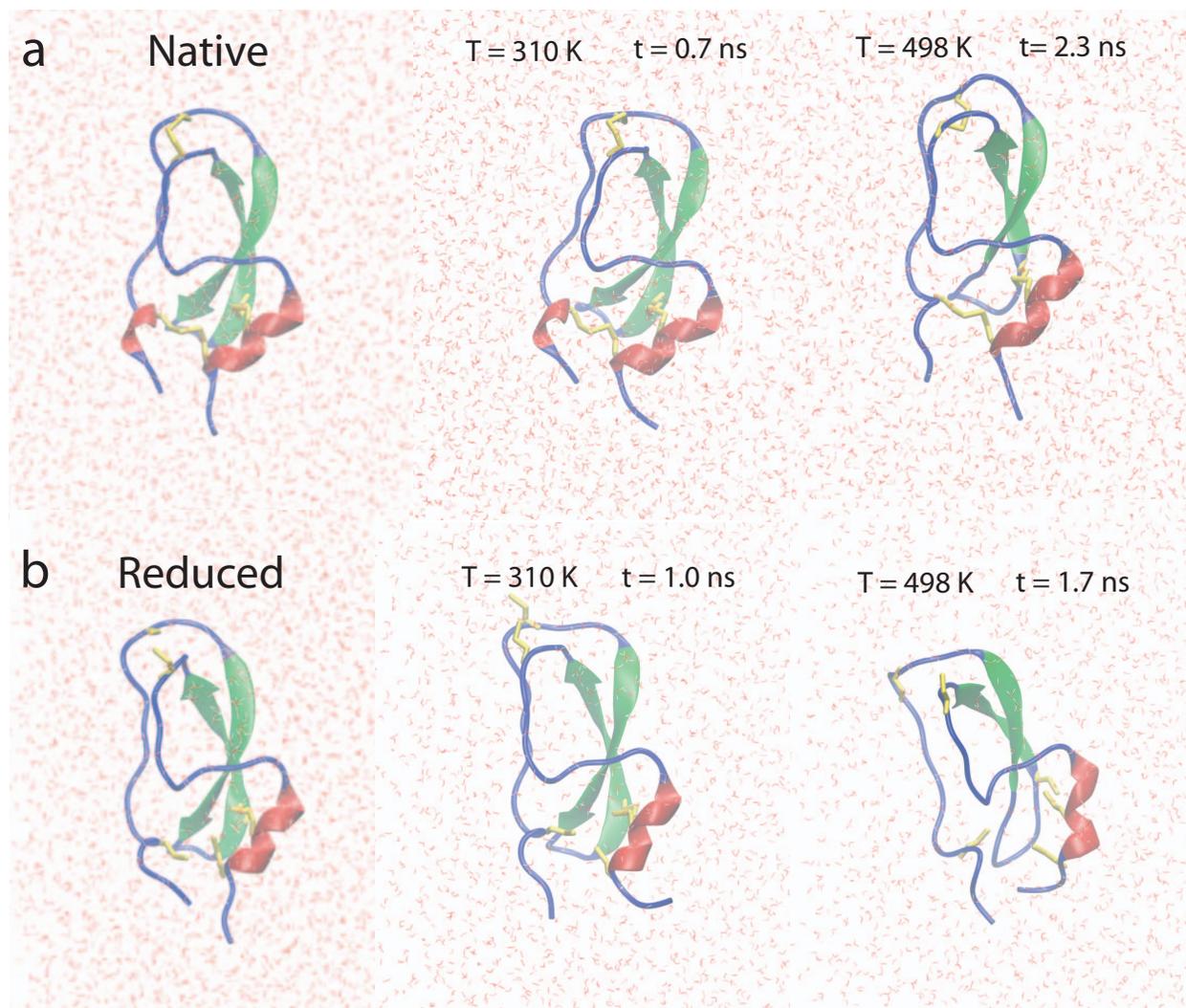


Figure 8: Structures of native and fully reduced BPTI. Shown are structural snapshots of the starting and final conformations of native and fully reduced BPTI at T=310K and T=498K. This figure is part of Exercise 5.

- Is reduced BPTI stable after being "boiled" at T=498K for t=1.7ns? How about the native one, is it also stable at T=498K after t=2.3 ns?
- Fig. 9 shows two contact maps of the native and the reduced BPTI system at T=498K after t=2.3ns and t=1.7 ns,

respectively. By comparing these two contact maps with each other and with that of the BPTI's crystal structure in Fig. 5, can you tell whether the two systems are stable at $T=498\text{K}$ and if not which parts are unfolding?

4 Trypsin-BPTI complex

This last section deals with BPTI and its interaction with other proteins in the cellular environment. The human body, e.g., needs a constant supply of amino acids for the repair and growth of its cells. The source of many of these amino acids are proteins that are ingested as parts of our diet. These proteins are then chopped up into their constituent amino acids by specialized enzymes in our digestive system. Trypsin is one such digestive enzyme [8] and is part of the digestive liquid that is produced by the pancreas.

Trypsin uses a special serine amino acid to cleave proteins specifically at a position near a lysine or arginine residue [8]. The cutting action of trypsin is so efficient that it has to be carefully controlled by inhibitor proteins such as BPTI, to prevent the degradation of arbitrary proteins, particularly in the very cells where trypsin is manufactured. BPTI binds to the serine site of trypsin that is responsible for the cleavage reaction and, thereby, prevents trypsin from performing its function. Once the trypsin-BPTI complex arrives in the digestive tract, BPTI is released and trypsin begins to digest. BPTI's ability to specifically inhibit trypsin is due to a special reactive site

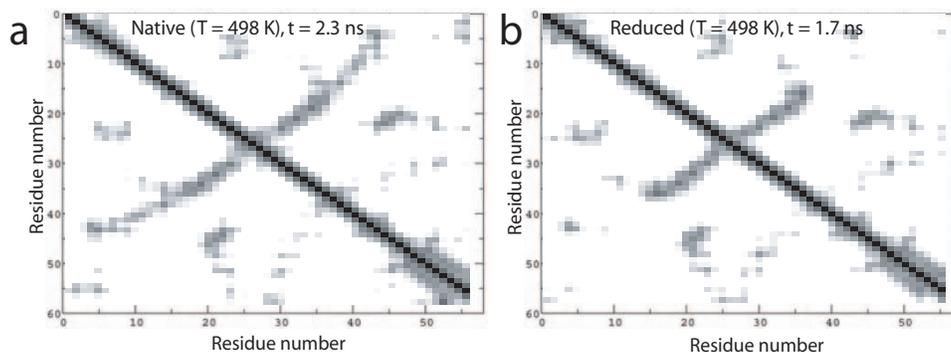


Figure 9: Contact maps of BPTI. Shown are the contact maps for the final frame of the native and fully reduced BPTI simulations at a temperature of $T=498\text{K}$. The figure is part of Exercise 5.

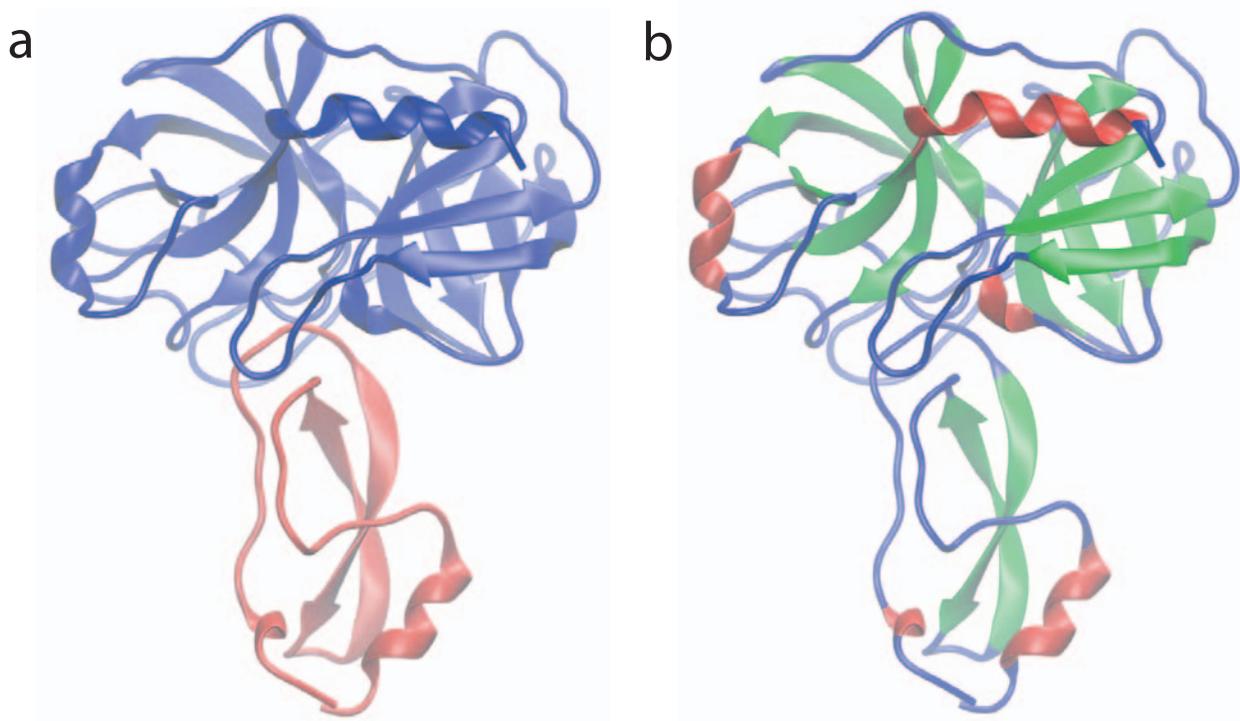


Figure 10: Trypsin-BPTI complex. Show are the trypsin-BPTI complex colored by segment BPTI (red) and trypsin (blue) (a), and the trypsin-BPTI complex colored by structure (b).

between residues Lys15 and Ala16 that can bind to trypsin without itself being cleaved.

Fig. 10 shows the crystal structure of the trypsin-BPTI complex found in the bovine pancreas [3]. The complex has a mushroom shape with trypsin forming the head of the mushroom and BPTI forming the stalk. The complex is stable and likely resembles other trypsin-protein complex intermediates formed during protein cleavage by trypsin.

BPTI contacts trypsin using a loop region located between residues Pro13 and Arg17. Upon docking to trypsin, this loop forms an anti parallel β -sheet structure with the cutting site of trypsin [1]. The cleavage activity of trypsin with proteins relies on the vicinity of the carbonyl carbon of the peptide bond to be cleaved to the sidechain carbonyl of the special serine residue, Ser195 [8]. To be more specific, trypsin uses a collection of three

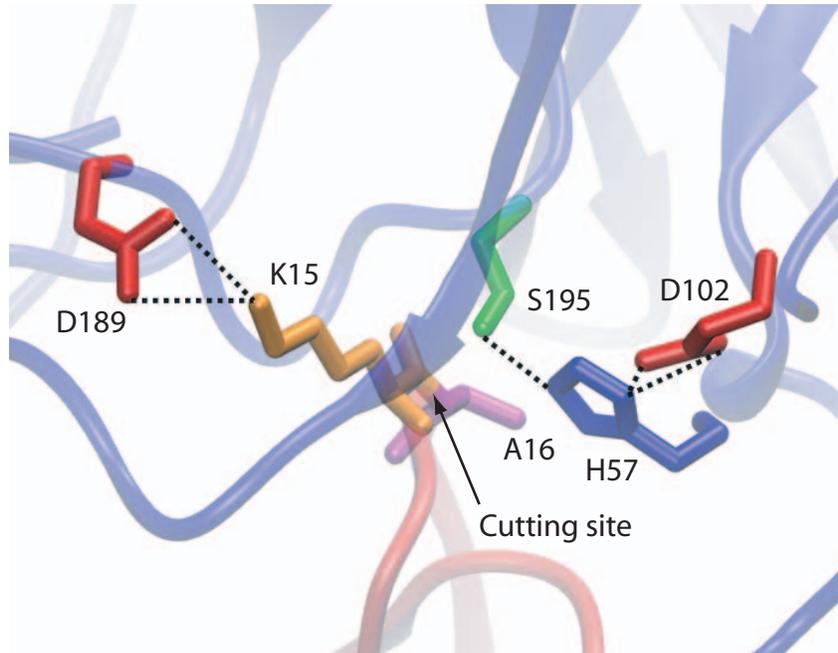


Figure 11: Cutting site of trypsin. Shown is the contact region between trypsin and BPTI involving residues Ser195, His57 and Asp102. These residues form a charge relay system and together with the negatively charged Asp189 aid in selectively cleaving a peptide bond near the lysine of a protein, BPTI in this case.

amino acid residues, Ser195, His57 and Asp102, that form a charge relay system to efficiently cleave the peptide bond. Furthermore, the negatively charged trypsin residue Asp189 ensures that proteins are selectively cut only near positively charged lysine or arginine residues due to its electrostatic interaction. The peptide bond on BPTI that interacts with the cleavage site on trypsin is located on the above mentioned loop region between residues Lys15 and Ala16. In contrast to other proteins, however, BPTI has evolved such that the peptide bond between Lys15 and Ala16 is held far enough from Ser195 to avoid cleavage, thereby, allowing it to act as a trypsin inhibitor. This is shown in Fig. 11 and will be explored in more detail in Exercise 6.

Exercise 6: Trypsin-BPTI complex

We will now take a closer look at the contact region between trypsin and BPTI where proteolytic cleavage normally occurs.

First, load the coordinate file `trypsin-bpti-2PTC.pdb` that contains the trypsin-BPTI complex into VMD and render the system such that it looks similar to the view shown in Fig. 10 (a).

- Can you identify the cutting site of trypsin (c.f. Fig. 11) at the contact region between trypsin and BPTI using VMD's sequence viewer or selection tool?
- By looking at the conformation of the charge relay system involving Ser195, His57 and Asp102 can you determine how this relay enhances the ability to cut the peptide bond?

References

- [1] Ascenzi, P., Bocedi, A., Bolognesi, M., Spallarossa, A., Coletta, M., De Cristofaro, R., Menegatti, E. (2003). The bovine basic pancreatic trypsin inhibitor (Kunitz inhibitor): a milestone protein., *Curr. Protein Peptide Sci.* **4**, 231-251.
- [2] Protein Data Bank, <http://www.rcsb.org/pdb/>
- [3] Marquart, M., Walter, J., Deisenhofer, J., Bode, W., Huber, R. (1983). The Geometry of the Reactive Site and of the Peptide Groups in Trypsin, Trypsinogen and its Complexes with Inhibitors *Acta Crystallogr., Sect. B* **39**, 480.
- [4] Stone, J., Wright, D., Eargle, J., Khalili, F., Villa, E., Tajkhorshid, E., Dhaliwal, B., Luthey-Schulten, Z. (Dec. 2004) Aquaporins with the VMD MultiSeq Tool (Tutorial), <http://www.ks.uiuc.edu/Training/Tutorials/>
- [5] Chen, C., Hsu, C. H., Su, N. Y., Lin, Y. C., Chiou, S. H., Wu, S. H. (2001). Solution Structure of a Kunitz-Type Chymotrypsin Inhibitor Isolated from the Elapid Snake Bungarus Fasciatus *J. Biol. Chem.* **276**, 45079-45087.
- [6] Antuch, W., Berndt, K. D., Chavez, M. A., Delfin, J., Wuthrich, K. (1993). The NMR solution structure of a Kunitz-type proteinase inhibitor from the sea anemone *Stichodactyla helianthus*. *Eur. J. Biochem.* **212**, 675-684.

- [7] Hynes, T. R., Randal, M., Kennedy, L. A., Eigenbrot, C., Kossiakoff, A. A. (1990). X-ray crystal structure of the protease inhibitor domain of Alzheimer's amyloid beta-protein precursor. *Biochemistry* **29**, 10018-10022.
- [8] Goodsell, D.S. (Oct. 2003). Molecule of the Month: Trypsin, http://www.rcsb.org/pdb/molecules/pdb46_1.html